

# SEQUENCE LISTING

<110> Darst, Seth A  
Zhang, Gongyi  
Campbell, ELizabeth  
Minakin, Leonid  
Severinov, Konstantin

<120> A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS  
OF USE THEREOF

<130> 600-1-258

<140> UNASSIGNED

<141> 1999-09-15

<160> 4

<170> PatentIn Ver. 2.0

<210> 1

<211> 1525

<212> PRT

<213> Thermus aquaticus

<220>

<221> SITE

<222> (1247)

<223> Any amino acid can be at this position

<400> 1

Met Lys Lys Glu Val Arg Lys Val Arg Ile Ala Leu Ala Ser Pro Glu  
1 5 10 15

Lys Ile Arg Ser Trp Ser Tyr Gly Glu Val Glu Lys Pro Glu Thr Ile  
20 25 30

Asn Tyr Arg Thr Leu Lys Pro Glu Arg Asp Gly Leu Phe Asp Glu Arg  
35 40 45

Ile Phe Gly Pro Ile Lys Asp Tyr Glu Cys Ala Cys Gly Lys Tyr Lys  
50 55 60

Arg Gln Arg Phe Glu Gly Lys Val Cys Glu Arg Cys Gly Val Glu Val  
65 70 75 80

Thr Arg Ser Ile Val Arg Arg Tyr Arg Met Gly His Ile Glu Leu Ala  
85 90 95

Thr Pro Ala Ala His Ile Trp Phe Val Lys Asp Val Pro Ser Lys Ile  
100 105 110

Gly Thr Leu Leu Asp Leu Phe Ala Thr Glu Leu Glu Gln Val Leu Tyr  
115 120 125

Phe Asn Lys Tyr Ile Val Leu Asp Pro Lys Gly Ala Val Leu Asp Gly  
130 135 140

Val Pro Val Glu Lys Arg Gln Leu Leu Thr Asp Glu Glu Tyr Arg Glu  
145 150 155 160

Leu Arg Tyr Gly Lys Gln Glu Thr Tyr Pro Leu Pro Ala Gly Val Asp  
165 170 175

Ala Leu Val Lys Asp Gly Glu Glu Val Val Lys Gly Gln Glu Leu Ala  
180 185 190

Pro Gly Val Val Ser Arg Met Asp Gly Val Gly Ser Leu Pro Leu Pro  
195 200 205

Arg Arg Val Arg Val Asp Tyr Leu Arg Lys Glu Arg Ala Ala Leu Arg  
210 215 220

Ile Pro Leu Ser Ala Trp Val Glu Lys Glu Pro Tyr Arg Pro Gly Glu  
225 230 235 240

Val Leu Ala Glu Leu Ser Glu Pro Tyr Leu Phe Arg Ala Glu Glu Ser  
245 250 255

Gly Val Val Glu Leu Lys Asp Leu Ala Glu Gly His Leu Ile Tyr Leu  
260 265 270

Arg Gln Glu Glu Glu Val Val Ala Arg Tyr Phe Leu Pro Ala Gly Met  
275 280 285

Thr Pro Leu Val Val Glu Gly Glu Ile Val Glu Val Gly Gln Pro Leu  
290 295 300

Ala Glu Gly Lys Gly Leu Leu Arg Leu Pro Arg His Met Thr Ala Lys  
305 310 315 320

Glu Val Glu Ala Glu Glu Gly Asp Ser Val His Leu Thr Leu Phe  
325 330 335

Leu Glu Trp Thr Glu Pro Lys Asp Tyr Lys Val Ala Pro His Met Asn  
340 345 350

Val Ile Val Pro Glu Gly Ala Lys Val Gln Ala Gly Glu Lys Ile Val  
355 360 365

Ala Ala Ile Asp Pro Glu Glu Glu Val Ile Ala Gln Ala Glu Gly Val  
370 375 380

Val His Leu His Glu Pro Ala Ser Ile Leu Val Val Lys Ala Arg Val  
385 390 395 400

Tyr Pro Phe Glu Asp Asp Val Glu Val Thr Thr Gly Asp Arg Val Ala  
405 410 415

Pro Gly Asp Val Leu Ala Asp Gly Gly Lys Val Lys Ser Glu Ile Tyr  
420 425 430

Gly Arg Val Glu Val Asp Leu Val Arg Asn Val Val Arg Val Val Glu  
435 440 445

Ser Tyr Asp Ile Asp Ala Arg Met Gly Ala Glu Ala Ile Gln Glu Leu  
450 455 460

Leu Lys Glu Leu Asp Leu Glu Lys Leu Glu Arg Glu Leu Leu Glu Glu  
465 470 475 480

Met Lys His Pro Ser Arg Ala Arg Arg Ala Lys Ala Arg Lys Arg Leu  
485 490 495

Glu Val Val Arg Ala Phe Leu Asp Ser Gly Asn Arg Pro Glu Trp Met  
500 505 510

Ile Leu Glu Ala Val Pro Val Leu Pro Pro Asp Leu Arg Pro Met Val  
515 520 525

Gln Val Asp Gly Gly Arg Phe Ala Thr Ser Asp Leu Asn Asp Leu Tyr  
530 535 540

Arg Arg Leu Ile Asn Arg Asn Asn Arg Leu Lys Lys Leu Leu Ala Gln  
545 550 555 560

Gly Ala Pro Glu Ile Ile Ile Arg Asn Glu Lys Arg Met Leu Gln Glu  
565 570 575

Ala Val Asp Ala Val Ile Asp Asn Gly Arg Arg Gly Ser Pro Val Thr  
580 585 590

Asn Pro Gly Ser Glu Arg Pro Leu Arg Ser Leu Thr Asp Ile Leu Ser  
595 600 605

Gly Lys Gln Gly Arg Phe Arg Gln Asn Leu Leu Gly Lys Arg Val Asp  
610 615 620

Tyr Ser Gly Arg Ser Val Ile Val Val Gly Pro Gln Leu Lys Leu His  
625 630 635 640

Gln Cys Gly Leu Pro Lys Arg Met Ala Leu Glu Leu Phe Lys Pro Phe  
645 650 655

Leu Leu Lys Lys Met Glu Glu Lys Ala Phe Ala Pro Asn Val Lys Ala  
660 665 670

Ala Arg Arg Met Leu Glu Arg Gln Arg Asp Ile Lys Asp Glu Val Trp  
675 680 685

Asp Ala Leu Glu Glu Val Ile His Gly Lys Val Val Leu Leu Asn Arg  
690 695 700

Ala Pro Thr Leu His Arg Leu Gly Ile Gln Ala Phe Gln Pro Val Leu  
705 710 715 720

Val Glu Gly Gln Ser Ile Gln Leu His Pro Leu Val Cys Glu Ala Phe  
725 730 735

Asn Ala Asp Phe Asp Gly Asp Gln Met Ala Val His Val Pro Leu Ser  
740 745 750

Ser Phe Ala Gln Ala Glu Ala Arg Ile Gln Met Leu Ser Ala His Asn  
755 760 765

Leu Leu Ser Pro Ala Ser Gly Glu Pro Leu Ala Lys Pro Ser Arg Asp  
770 775 780

Ile Ile Leu Gly Leu Tyr Tyr Ile Thr Gln Val Arg Lys Glu Lys Lys  
785 790 795 800

Gly Ala Gly Met Ala Phe Ala Thr Pro Glu Glu Ala Leu Ala Ala Tyr  
805 810 815

Glu Arg Gly Glu Val Ala Leu Asn Ala Pro Ile Val Val Ala Gly Arg  
820 825 830

Glu Thr Ser Val Gly Arg Leu Lys Phe Val Phe Ala Asn Pro Asp Glu  
835 840 845

Ala Leu Leu Ala Val Ala His Gly Leu Leu Asp Leu Gln Asp Val Val  
850 855 860

Thr Val Arg Tyr Leu Gly Arg Arg Leu Glu Thr Asn Pro Gly Arg Ile  
865 870 875 880

Leu Phe Ala Arg Ile Val Gly Glu Ala Val Gly Asp Glu Lys Val Ala  
885 890 895

Gln Glu Leu Ile Gln Met Asp Val Pro Gln Glu Lys Asn Ser Leu Lys  
900 905 910

Asp Leu Val Tyr Gln Ala Phe Leu Arg Leu Gly Met Glu Lys Thr Ala  
915 920 925

Arg Leu Leu Asp Ala Leu Lys Tyr Tyr Gly Phe Thr Leu Ser Thr Thr  
930 935 940

Ser Gly Ile Ile Thr Ile Gly Ile Asp Asp Ala Val Ile Pro Glu Glu  
945 950 955 960

Lys Gln Arg Tyr Leu Glu Glu Ala Asp Arg Lys Leu Arg Gln Ile Glu  
965 970 975

Gln Ala Tyr Glu Met Gly Phe Leu Thr Asp Arg Glu Arg Tyr Asp Gln  
980 985 990

Val Ile Gln Leu Trp Thr Glu Thr Thr Glu Lys Val Thr Gln Ala Val  
995 1000 1005

Phe Asn Asn Phe Glu Glu Asn Tyr Pro Phe Asn Pro Leu Tyr Val Met  
1010 1015 1020

Ala Gln Ser Gly Ala Arg Gly Asn Pro Gln Gln Ile Arg Gln Leu Cys  
1025 1030 1035 1040

Gly Met Arg Gly Leu Met Gln Lys Pro Ser Gly Glu Thr Phe Glu Val  
1045 1050 1055

Pro Val Arg Ser Ser Phe Arg Glu Gly Leu Thr Val Leu Glu Tyr Phe  
1060 1065 1070

Ile Ser Ser His Gly Ala Arg Lys Gly Gly Ala Asp Thr Ala Leu Arg  
1075 1080 1085

Thr Ala Asp Ser Gly Tyr Leu Thr Arg Lys Leu Val Asp Val Ala His  
1090 1095 1100

Glu Ile Val Val Arg Glu Ala Asp Cys Gly Thr Thr Lys Tyr Ile Ser  
1105 1110 1115 1120

Val Pro Leu Phe Gln Met Asp Glu Val Thr Arg Thr Leu Arg Leu Arg  
1125 1130 1135

Lys Arg Ser Asp Ile Glu Ser Gly Leu Tyr Gly Arg Val Leu Ala Arg  
1140 1145 1150

Glu Val Glu Ala Leu Gly Arg Arg Leu Glu Glu Gly Arg Tyr Leu Ser  
1155 1160 1165

Leu Glu Asp Val His Phe Leu Ile Lys Ala Ala Glu Ala Gly Glu Val  
1170 1175 1180

Arg Glu Val Pro Val Arg Ser Pro Leu Thr Cys Gln Thr Arg Tyr Gly  
1185 1190 1195 1200

Val Cys Gln Lys Cys Tyr Gly Tyr Asp Leu Ser Met Ala Arg Pro Val  
1205 1210 1215

Ser Ile Gly Glu Ala Val Gly Val Val Ala Ala Glu Ser Ile Gly Glu  
1220 1225 1230

Pro Gly Thr Gln Leu Thr Met Arg Thr Phe His Thr Gly Gly Xaa Ala  
1235 1240 1245

Val Gly Thr Asp Ile Thr Gln Gly Leu Pro Arg Val Ile Glu Leu Phe  
1250 1255 1260

Glu Ala Arg Arg Pro Lys Ala Lys Ala Val Ile Ser Glu Ile Asp Gly  
1265 1270 1275 1280

Val Val Arg Ile Glu Glu Gly Glu Asp Arg Leu Ser Val Phe Val Glu  
1285 1290 1295

Ser Glu Gly Phe Ser Lys Glu Tyr Lys Leu Pro Lys Asp Ala Arg Leu  
1300 1305 1310

Leu Val Lys Asp Gly Asp Tyr Val Glu Ala Gly Gln Pro Leu Thr Arg  
1315 1320 1325

Gly Ala Ile Asp Pro His Gln Leu Leu Glu Ala Lys Gly Pro Glu Ala  
1330 1335 1340

Val Glu Arg Tyr Leu Val Asp Glu Ile Gln Lys Val Tyr Arg Ala Gln  
1345 1350 1355 1360

Gly Val Lys Leu His Asp Lys His Ile Glu Ile Val Val Arg Gln Met  
1365 1370 1375



Asp Val Pro Pro Glu Lys Arg Glu Asn Val Gly Ile Gln Ala Ala Phe  
35 40 45

Lys Glu Thr Phe Pro Ile Glu Glu Gly Asp Lys Gly Lys Gly Gly Leu  
50 55 60

Val Leu Asp Phe Leu Glu Tyr Arg Ile Gly Asp Pro Pro Phe Ser Gln  
65 70 75 80

Asp Glu Cys Arg Glu Lys Asp Leu Thr Tyr Gln Ala Pro Leu Tyr Ala  
85 90 95

Arg Leu Gln Leu Ile His Lys Asp Thr Gly Leu Ile Lys Glu Asp Glu  
100 105 110

Val Phe Leu Gly His Leu Pro Leu Met Thr Glu Asp Gly Ser Phe Ile  
115 120 125

Ile Asn Gly Ala Asp Arg Val Ile Val Ser Gln Ile His Arg Ser Pro  
130 135 140

Gly Val Tyr Phe Thr Pro Asp Pro Ala Arg Pro Gly Arg Tyr Ile Ala  
145 150 155 160

Ser Ile Ile Pro Leu Pro Lys Arg Gly Pro Trp Ile Asp Leu Glu Val  
165 170 175

Glu Ala Ser Gly Val Val Thr Met Lys Val Asn Lys Arg Lys Phe Pro  
180 185 190

Leu Val Leu Leu Leu Arg Val Leu Gly Tyr Asp Gln Glu Thr Leu Val  
195 200 205

Arg Glu Leu Ser Ala Tyr Gly Asp Leu Val Gln Gly Leu Leu Asp Glu  
210 215 220

Ala Val Leu Ala Met Arg Pro Glu Glu Ala Met Val Arg Leu Phe Thr  
225 230 235 240

Leu Leu Arg Pro Gly Asp Pro Pro Lys Lys Asp Lys Ala Leu Ala Tyr  
245 250 255

Leu Phe Gly Leu Leu Ala Asp Pro Lys Arg Tyr Asp Leu Gly Glu Ala  
260 265 270

Gly Arg Tyr Lys Ala Glu Glu Lys Leu Gly Val Gly Leu Ser Gly Arg  
275 280 285



Thr Leu Val Arg Phe Glu Asp Gly Glu Phe Lys Asp Glu Val Phe Leu  
290 295 300

Pro Thr Leu Arg Tyr Leu Phe Ala Leu Thr Ala Gly Val Pro Gly His  
305 310 315 320

Glu Val Asp Asp Ile Asp His Leu Gly Asn Arg Arg Ile Arg Thr Val  
325 330 335

Gly Glu Leu Met Ala Asp Gln Phe Arg Val Gly Leu Ala Arg Leu Ala  
340 345 350

Arg Gly Val Arg Glu Arg Met Val Met Gly Ser Pro Asp Thr Leu Thr  
355 360 365

Pro Ala Lys Leu Val Asn Ser Arg Pro Leu Glu Ala Ala Leu Arg Glu  
370 375 380

Phe Phe Ser Arg Ser Gln Leu Ser Gln Phe Lys Asp Glu Thr Asn Pro  
385 390 395 400

Leu Ser Ser Leu Arg His Lys Arg Arg Ile Ser Ala Leu Gly Pro Gly  
405 410 415

Gly Leu Thr Arg Glu Arg Ala Gly Phe Asp Val Arg Asp Val His Arg  
420 425 430

Thr His Tyr Gly Arg Ile Cys Pro Val Glu Thr Pro Glu Gly Ala Asn  
435 440 445

Ile Gly Leu Ile Thr Ser Leu Ala Ala Tyr Ala Arg Val Asp Ala Leu  
450 455 460

Gly Phe Ile Arg Thr Pro Tyr Arg Arg Val Lys Asn Gly Val Val Thr  
465 470 475 480

Glu Glu Val Val Tyr Met Thr Ala Ser Glu Glu Asp Arg Tyr Thr Ile  
485 490 495

Ala Gln Ala Asn Thr Pro Leu Glu Gly Asp Arg Ile Ala Thr Asp Arg  
500 505 510

Val Val Ala Arg Arg Arg Gly Glu Pro Val Ile Val Ala Pro Glu Glu  
515 520 525

Val Glu Phe Met Asp Val Ser Pro Lys Gln Val Phe Ser Leu Asn Thr  
530 535 540

Asn Leu Ile Pro Phe Leu Glu His Asp Asp Ala Asn Arg Ala Leu Met  
 545 550 555 560  
 Gly Ser Asn Met Gln Thr Gln Ala Val Pro Leu Ile Arg Ala Gln Ala  
 565 570 575  
 Pro Val Val Met Thr Gly Leu Glu Glu Arg Val Val Arg Asp Ser Leu  
 580 585 590  
 Ala Ala Leu Tyr Ala Glu Glu Asp Gly Glu Val Val Lys Val Asp Gly  
 595 600 605  
 Thr Arg Ile Ala Val Arg Tyr Glu Asp Gly Arg Leu Val Glu His Pro  
 610 615 620  
 Leu Arg Arg Tyr Ala Arg Ser Asn Gln Gly Thr Ala Phe Asp Gln Arg  
 625 630 635 640  
 Pro Arg Val Arg Val Gly Gln Arg Val Lys Lys Gly Asp Leu Leu Ala  
 645 650 655  
 Asp Gly Pro Ala Ser Glu Glu Gly Phe Leu Ala Leu Gly Gln Asn Val  
 660 665 670  
 Leu Val Ala Ile Met Pro Phe Asp Gly Tyr Asn Phe Glu Asp Ala Ile  
 675 680 685  
 Val Ile Ser Glu Glu Leu Xaa Xaa Arg Asp Phe Tyr Thr Ser Ile His  
 690 695 700  
 Ile Glu Arg Tyr Glu Ile Glu Ala Arg Asp Thr Lys Leu Gly Pro Glu  
 705 710 715 720  
 Arg Ile Thr Arg Asp Ile Pro His Leu Ser Glu Ala Ala Leu Arg Asp  
 725 730 735  
 Leu Asp Glu Glu Gly Ile Val Arg Ile Gly Ala Glu Val Lys Pro Gly  
 740 745 750  
 Asp Ile Leu Val Gly Arg Thr Ser Phe Lys Gly Glu Gln Glu Pro Ser  
 755 760 765  
 Pro Glu Glu Arg Leu Leu Arg Ser Ile Phe Gly Glu Lys Ala Arg Asp  
 770 775 780  
 Val Lys Asp Thr Ser Leu Arg Val Pro Pro Gly Glu Gly Gly Ile Val  
 785 790 795 800

Val Gly Arg Leu Arg Leu Arg Arg Gly Asp Pro Gly Val Glu Leu Lys  
805 810 815

Pro Gly Val Arg Glu Val Val Arg Val Phe Val Ala Gln Lys Arg Lys  
820 825 830

Leu Gln Val Gly Asp Lys Leu Ala Asn Arg His Gly Asn Lys Gly Val  
835 840 845

Val Ala Lys Ile Leu Pro Val Glu Asp Met Pro His Leu Pro Asp Gly  
850 855 860

Thr Pro Val Asp Val Ile Leu Asn Pro Leu Gly Val Pro Ser Arg Met  
865 870 875 880

Asn Leu Gly Gln Ile Leu Glu Thr His Leu Gly Leu Ala Gly Tyr Phe  
885 890 895

Leu Gly Gln Arg Tyr Ile Ser Pro Val Phe Asp Gly Ala Thr Glu Pro  
900 905 910

Glu Ile Lys Glu Leu Leu Ala Glu Ala Phe Asn Leu Tyr Phe Gly Lys  
915 920 925

Arg Gln Gly Glu Gly Phe Gly Val Asp Lys Arg Glu Lys Glu Val Leu  
930 935 940

Ala Arg Ala Glu Lys Leu Gly Leu Val Ser Pro Gly Lys Ser Pro Glu  
945 950 955 960

Glu Gln Leu Lys Glu Leu Phe Asp Leu Gly Lys Val Val Leu Tyr Asp  
965 970 975

Gly Arg Thr Gly Glu Pro Phe Glu Gly Pro Ile Val Val Gly Gln Met  
980 985 990

Phe Ile Met Lys Leu Tyr His Met Val Glu Asp Lys Met His Ala Arg  
995 1000 1005

Ser Thr Gly Pro Tyr Ser Leu Ile Thr Gln Gln Pro Leu Gly Gly Lys  
1010 1015 1020

Ala Gln Phe Gly Gly Gln Arg Phe Gly Glu Met Glu Val Trp Ala Leu  
1025 1030 1035 1040

Glu Ala Tyr Gly Ala Ala His Thr Leu Gln Glu Met Leu Thr Ile Lys  
1045 1050 1055



145

150

155

160

Ile Asn Ala Ile Pro Val Asp Ala Ile Phe Ser Pro Val Arg Arg Val  
 165 170 175

Ala Phe Gln Val Glu Asp Thr Arg Leu Gly Gln Arg Thr Asp Leu Asp  
 180 185 190

Lys Leu Thr Leu Arg Ile Trp Thr Asp Gly Ser Val Thr Pro Leu Glu  
 195 200 205

Ala Leu Asn Gln Ala Val Ala Ile Leu Lys Glu His Leu Asn Tyr Phe  
 210 215 220

Ala Asn Pro Glu Ala Ser Leu Leu Pro Thr Pro Glu Val Ser Lys Gly  
 225 230 235 240

Glu Lys Arg Glu Ser Ala Glu Glu Asp Leu Asp Leu Pro Leu Glu Glu  
 245 250 255

Leu Gly Leu Ser Thr Arg Val Leu His Ser Leu Lys Glu Glu Gly Ile  
 260 265 270

Glu Ser Val Arg Ala Leu Leu Ala Leu Asn Leu Lys Asp Leu Arg Asn  
 275 280 285

Ile Pro Gly Ile Gly Glu Arg Ser Leu Glu Glu Ile Arg Gln Ala Leu  
 290 295 300

Ala Lys Lys Gly Phe Thr Leu Lys Glu  
 305 310

&lt;210&gt; 4

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: natural part  
 of bacterial proteins

&lt;400&gt; 4

Asn Ala Asp Phe Asp Gly Asp

1

5